
7E

Action

Professional Services Committee of the Whole

Adoption of Passing Score Standard for the California Subject Matter Examinations for Teachers (CSET): Filipino Examination

Executive Summary: This report provides the Commission with recommendations relevant to the determination of the passing standard for the CSET: Filipino examination.

Recommended Action: That the Commission adopt the recommended passing standard for the CSET: Filipino examination.

Presenter: Phyllis Jacobson, Administrator, Professional Services Division, and Marty Karlin, Senior Area Director, National Evaluation Systems, Inc.

Strategic Plan Goal: 1

Promote educational excellence through the preparation and certification of professional educators

- ◆ Sustain high quality standards for the preparation of professional educators.
- ◆ Assess and monitor the efficacy of the Accreditation System, Examination System, and State and Federal Funded Programs.

November-December 2006

Adoption of Passing Score Standard for the California Subject Matter Examinations for Teachers (CSET): Filipino Examination

Introduction

This report describes the standard setting study for the California Subject Examinations for Teachers (CSET): Languages Other Than English in Filipino, and provides staff-recommended initial passing standards for the examination.

Background

In summer 2006, the Commission appointed a subject matter advisory panel for Phase IV of the California Subject Examinations for Teachers (CSET) for the single subject area of Filipino to advise Commission staff on the development of an examination in this subject area. National Evaluation Systems, Inc. (NES), the Commission's CSET testing contractor, and Commission staff have worked with the panel since then to facilitate this work. The subject matter advisory panel consisted of classroom teachers, subject area specialists, college and university faculty, and members of relevant professional organizations and committees, all with a specialty in the subject area.

From spring through fall of 2003, an advisory panel of educators developed the Subject Matter Requirements (SMRs) for Languages Other Than English, which were aligned with state and national student content standards and frameworks, and standards of national professional organizations. The Commission approved the SMRs for Languages Other Than English at its January 2004 meeting. The SMRs specify the content that is to be taught in Commission-approved subject matter preparation programs and that is eligible for assessment on the Commission's subject matter examinations.

The panel also developed new program standards for Languages Other Than English based upon the content in the SMRs that will be utilized by California accredited colleges and universities to develop single subject matter preparation programs in these academic areas. Those program standards were approved by the Commission at its May 2004 meeting.

The development of the Filipino examination occurred during the fall of 2006. A test structure was developed based on the previously approved test structures for Languages Other Than English, and constructed-response items were drafted, reviewed, and revised as needed by both the Bias Review Committee and the subject matter advisory panel. The subject matter advisory panel selected marker responses and scored the responses from the first administration of the examination, which occurred on November 4, 2006. Additionally, a test guide including the subject matter requirements, test structure, and sample questions was developed to assist candidates in preparing to take the new examination.

On November 4, 2006, the first test administration of the new CSET: Filipino was held. On November 15, 2006, the standard setting study for this examination was held in Sacramento to determine the initial passing standard recommendations of California educators.

The CSET: Filipino Examination

The CSET: Filipino examination is a paper-and-pencil test that is divided into two subtests, each made up of six or seven constructed-response items. Constructed-response items are of two types: *extended* constructed-response items that are scored using a four-point scale, and *focused* constructed-response items that are scored using a three-point scoring scale. The CSET Filipino examination requires oral responses spoken in Filipino to some test items. The test structure for CSET: Filipino is shown in Table 1 of this agenda item. Constructed-response performance characteristics and scoring scales are provided in Appendix A.

Each CSET testing session is five hours in length. Examinees can choose to take either or both subtests within a single testing session. Individual subtests are not timed. The CSET: Filipino will be administered two times each year. The numbers of examinees who took each subtest at the first test administration of the test in November 2006 are provided in Appendix B. On November 15, 2006, Commission staff and NES conducted a standard setting study for the new examination. The standard-setting procedures used and the results of the study are described later in this report.

Table 1: Test Structure of the CSET: Filipino

Subtest		Number of Items
I	<ul style="list-style-type: none"> • General Linguistics • Linguistics of the Target Language • Literary and Cultural Texts and Traditions • Cultural Analysis and Comparisons 	1 focused 3 focused 1 extended 1 extended
	Total Subtest I	4 focused 2 extended
II	<ul style="list-style-type: none"> • Listening Comprehension • Reading Comprehension • Written Expression • Oral Expression 	2 focused 2 focused 1 extended 2 focused
	Total Subtest II	6 focused 1 extended

The Standard Setting Study

A standard setting study for the CSET: Filipino was conducted on November 15, 2006 by a panel of educators with expertise in the subject area. The purpose of the standard setting study is to provide the Commission with recommendations, based on the informed judgments of California

educators, relevant to the determination of the initial passing standards for the CSET: Filipino. The number and the demographics of the panel members, including classroom teachers and college and university faculty, who participated in the study will be provided in an agenda insert for the November-December 2006 meeting.

The standard setting study began with an orientation and training session. Panel members were provided the subject matter requirements as well as the subtest forms used for the November 2006 test administration. To help the panel members become familiar with the examination, the knowledge and skills associated with the items, and the perspective of the examinee, panel members were asked to take the test under simulated test-like conditions. They were asked to read and answer each item independently.

Panel members were then asked to consider the “just acceptable” candidate. Although many of the examinees will exceed the level of knowledge and skills of the acceptably qualified candidate, none should fall below that level. For this reason, panel members were trained to make judgments based on candidates just at the level of knowledge and skills required of an entry-level teacher candidate to successfully satisfy the subject matter requirements.

After extensive training and the simulated test taking, panel members were asked to complete three rounds of standard setting tasks based on the test structure. This process is briefly described below. A more detailed description of the process is found in Appendix B.

In Round One, panel members were asked to individually rate each item on each subtest. They were asked to rate the level of response that would be achieved by the “just acceptable” candidate for each of the items on each subtest.

Using the item statistics produced from Round One to inform judgments, Round Two moved the panel from individual item ratings to ratings at the subtest level. Panelists were asked to provide, for each subtest, their estimation of the total score points that would be achieved by the “just acceptable” candidate on the entire set of items.

In the final round of ratings, the panel members were asked to make independent recommendations for a passing standard. To aid in their discussions, they were provided the results of the subtest-level statistics generated from Round Two, any applicable examinee demographic information, and data analyses on the percent of examinees from the first test administration who would pass each subtest, given all possible passing scores.

Results

The results of the standard setting study will be provided in an agenda insert at the November-December 2006 meeting.

APPENDIX A

PERFORMANCE CHARACTERISTICS,

SCORE SCALES,

AND

RELEVANT SUBJECT MATTER REQUIREMENTS

SCORING RUBRIC FOR FOCUSED-RESPONSE ITEMS

THREE-POINT SCORE SCALE

PERFORMANCE CHARACTERISTICS

PURPOSE	The extent to which the examinee responds to the constructed-response assignment's charge in relation to relevant CSET subject matter requirements.
SUBJECT MATTER KNOWLEDGE	The application of accurate subject matter knowledge as described in the relevant CSET subject matter requirements.
SUPPORT	The appropriateness and quality of the supporting evidence in relation to relevant CSET subject matter requirements.

SCORE SCALE

SCORE POINT	SCORE POINT DESCRIPTION
3	The "3" response reflects a command of the relevant knowledge and skills as defined in the CSET Subject Matter Requirements. <ul style="list-style-type: none">• The purpose of the assignment is fully achieved.• There is an accurate application of relevant subject matter knowledge.• There is appropriate and specific relevant supporting evidence.
2	The "2" response reflects a general command of the relevant knowledge and skills as defined in the CSET Subject Matter Requirements. <ul style="list-style-type: none">• The purpose of the assignment is largely achieved.• There is a largely accurate application of relevant subject matter knowledge.• There is acceptable relevant supporting evidence.
1	The "1" response reflects a limited or no command of the relevant knowledge and skills as defined in the CSET Subject Matter Requirements. <ul style="list-style-type: none">• The purpose of the assignment is only partially or not achieved.• There is limited or no application of relevant subject matter knowledge.• There is little or no relevant supporting evidence.

For Listening and Reading Comprehension Assignments:

U	The "U" (Unscorable) is assigned to a response that is unrelated to the assignment, illegible, not in the target language or English, or does not contain a sufficient amount of original work to score.
B	The "B" (Blank) is assigned to a response that is blank.

For Oral Expression Assignments:

U	The "U" (Unscorable) is assigned to a response that is unrelated to the assignment, inaudible/incomprehensible, not in the target language, or does not contain a sufficient amount of original work to score.
B	The "B" (Blank) is assigned to a response that is blank.

For all other assignments:

U	The "U" (Unscorable) is assigned to a response that is unrelated to the assignment, illegible, not in the target language or English, or does not contain a sufficient amount of original work to score.
B	The "B" (Blank) is assigned to a response that is blank.

RELEVANT SUBJECT MATTER REQUIREMENTS

Relevant Subject Matter Requirements for General Linguistics

- Demonstrate an understanding of the nature, purposes, and uses of language.
- Demonstrate an understanding of the development of language and the significance of language change, including the variations that occur within the contexts of time, place, age, gender, and situation.
- Demonstrate an understanding of the principles of pragmatics, discourse analysis, and the theory of speech acts.
- Demonstrate an understanding of theories of language acquisition and learning.

***Relevant Subject Matter Requirements for Linguistics of the Target Language
(Language Structures and Contrastive Analysis)***

- Demonstrate an understanding of the phonology of the target language.
- Demonstrate an understanding of the orthography of the target language.
- Demonstrate an understanding of the morphology of the target language.
- Demonstrate an understanding of the syntax of the target language.
- Demonstrate an understanding of the semantics of the target language.
- Describe changes that occur in the target language over time.
- Analyze and contrast linguistic structures of the target language and English.
- Compare and contrast particular words, idioms, and inflections in the target language and English.

***Relevant Subject Matter Requirements for Linguistics of the Target Language
(Sociolinguistics and Pragmatic)***

- Demonstrate an understanding of pragmatic and sociolinguistic features of target-language discourse.
- Demonstrate an understanding of the origins and social implications of accentual and dialectal differences within the target language.
- Describe the differences among the varieties of the target language and the factors that account for these differences.

Relevant Subject Matter Requirements for Listening Comprehension

- Demonstrate an understanding of the main ideas and significant details of oral messages in a variety of authentic contexts.
- Make deductive and inductive inferences based on information contained in oral messages.
- Analyze and evaluate oral messages in relation to their purpose, context, and point of view.

Relevant Subject Matter Requirements for Reading Comprehension

- Demonstrate an understanding of the main ideas and significant details of written texts.
- Make deductive and inductive inferences based on information contained in written texts.
- Apply critical reasoning skills to written texts.

Relevant Subject Matter Requirements for Oral Expression

- Construct connected oral discourse that communicates a message effectively.
- Construct connected oral discourse that demonstrates a wide range of vocabulary.
- Construct connected oral discourse that demonstrates idiomatic expressions.
- Construct connected oral discourse that demonstrates linguistic structures.

SCORING RUBRIC FOR TRANSFORMATION AND ERROR ANALYSIS ITEMS

PERFORMANCE CHARACTERISTICS

SUBJECT MATTER KNOWLEDGE	The application of accurate subject matter knowledge as described in the relevant CSET subject matter requirements.
---------------------------------	---

RELEVANT SUBJECT MATTER REQUIREMENTS

- Demonstrate an understanding of the orthography of the target language.
- Demonstrate an understanding of the morphology of the target language.
- Demonstrate an understanding of the syntax of the target language.
- Demonstrate an understanding of the semantics of the target language.
- Identify, analyze, and correct grammatical and mechanical errors in the target language.

Transformation

For these tasks, a “Correct” response is described by the following:

The candidate's response correctly transforms the sentence as instructed, i.e., the response follows all applicable grammatical rules of the target language and accurately supplies a transformed word, phrase, or clause construction as directed.

Error Analysis

For these tasks, a “Correct” response is described by the following:

The candidate's response acceptably corrects the syntactic or linguistic error(s) found in the sentence, i.e., the response has located and corrected the error(s) of language structure found within the sentence and has retained the original meaning.

SCORING RUBRIC FOR EXTENDED-RESPONSE ITEMS

FOUR-POINT SCORE SCALE

PERFORMANCE CHARACTERISTICS

PURPOSE	The extent to which the response addresses the constructed -response assignment's charge in relation to relevant CSET subject matter requirements.
SUBJECT MATTER KNOWLEDGE	The application of accurate subject matter knowledge as described in the relevant CSET subject matter requirements.
SUPPORT	The appropriateness and quality of the supporting evidence in relation to relevant CSET subject matter requirements.
DEPTH AND BREADTH OF UNDERSTANDING	The degree to which the response demonstrates understanding of the relevant CSET subject matter requirements.

SCORE SCALE

SCORE POINT	SCORE POINT DESCRIPTION
4	<p>The "4" response reflects a thorough command of the relevant knowledge and skills as defined in the CSET Subject Matter Requirements.</p> <ul style="list-style-type: none"> • The purpose of the assignment is fully achieved. • There is a substantial and accurate application of relevant subject matter knowledge. • The supporting evidence is sound; there are high-quality, relevant examples. • The response reflects a comprehensive understanding of the assignment.
3	<p>The "3" response reflects a general command of the relevant knowledge and skills as defined in the CSET Subject Matter Requirements.</p> <ul style="list-style-type: none"> • The purpose of the assignment is largely achieved. • There is a largely accurate application of relevant subject matter knowledge. • The supporting evidence is adequate; there are some acceptable, relevant examples. • The response reflects an adequate understanding of the assignment.
2	<p>The "2" response reflects a limited command of the relevant knowledge and skills as defined in the CSET Subject Matter Requirements.</p> <ul style="list-style-type: none"> • The purpose of the assignment is partially achieved. • There is limited accurate application of relevant subject matter knowledge. • The supporting evidence is limited; there are few relevant examples. • The response reflects a limited understanding of the assignment.
1	<p>The "1" response reflects little or no command of the relevant knowledge and skills as defined in the CSET Subject Matter Requirements.</p> <ul style="list-style-type: none"> • The purpose of the assignment is not achieved. • There is little or no accurate application of relevant subject matter knowledge. • The supporting evidence is weak; there are no or few relevant examples. • The response reflects little or no understanding of the assignment.

For Written Expression Assignments:

U	The "U" (Unscorable) is assigned to a response that is unrelated to the assignment, illegible, not in the target language, or does not contain a sufficient amount of original work to score.
B	The "B" (Blank) is assigned to a response that is blank.

For all other assignments:

U	The "U" (Unscorable) is assigned to a response that is unrelated to the assignment, illegible, not in the target language or English, or does not contain a sufficient amount of original work to score.
B	The "B" (Blank) is assigned to a response that is blank.

Relevant Subject Matter Requirements

Relevant Subject Matter Requirements for Literary and Cultural Texts and Traditions

- Demonstrate an understanding of major movements, genres, writers, and works in the literature of the target language.
- Demonstrate an understanding of the historical, social, and cultural influences on works of literature in the target language.
- Use knowledge of the literary and cultural traditions of the target culture to interpret changes in that culture over time.
- Demonstrate an understanding of the ways in which literary and intellectual works and movements of cultures associated with the target language both reflected and shaped those cultures.
 - Analyze and interpret a wide range of literary and cultural texts.
 - Evaluate the use of language to convey meaning, to inform, to persuade, or to evoke reader response.
 - Analyze the elements of literary works.
- Interpret the use of rhetorical and literary techniques.

Relevant Subject Matter Requirements for Cultural Analysis and Comparisons

- Demonstrate an understanding of how all of the cultural perspectives within nations and cultures associated with the target language interact to influence the development and evolution of the target cultures.
- Demonstrate familiarity with how the major physical and other geographical features of countries and cultures associated with the target language have influenced the cultures' development and evolution.
- Analyze how political factors have influenced the development and evolution of cultures associated with the target language, including the interrelationship between geography and political systems.
- Demonstrate familiarity with significant individuals, key eras, and major historical events and developments within nations and cultures associated with the target language, and analyze their influence on the development and evolution of the target cultures.
- Demonstrate an understanding of how the political, religious, social, economic, and educational systems and institutions in nations and cultures associated with the target language have been shaped by and have influenced the development and evolution of the target cultures.
- Demonstrate an understanding of how cultural practices exemplify cultural perspectives.
- Analyze cultural stereotypes and their effects on the perceptions of and attitudes toward the target cultures.
- Demonstrate an understanding of how the products of a target culture exemplify cultural perspectives.

Relevant Subject Matter Requirements for Written Expression

- Compose a well-organized passage in the target language employing styles and levels of diction appropriate for a given audience.
- Compose a well-organized passage in the target language employing styles and levels of diction appropriate for a given purpose.
- Compose a well-organized passage in the target language employing styles and levels of diction appropriate for a given occasion.
- Compose a well-organized passage in the target language demonstrating command of a wide range of vocabulary.
- Compose a well-organized passage in the target language demonstrating command of idiomatic expressions.
- Compose a well-organized passage in the target language demonstrating command of linguistic structures.

APPENDIX B

DETAILED STANDARD SETTING PROCESS

Standard Setting Rating Tasks

Round One Standard Setting Ratings

For Round One of the item ratings, panel members were provided the following materials.

- the subject matter requirements and subtest descriptions
- the subtest forms used for the November 4, 2006 test administration
- the appropriate set of performance characteristics and scoring scales
- the subject matter advisory panel-approved marker responses for each score point on the scoring scale, for each item
- the Round One Rating Form

To begin the Round One item ratings, panel members rated a practice set of sample items. They were asked to rate each item by responding to the following question.

:

Imagine a hypothetical candidate who is just at the level of knowledge and skills important for effective job performance as a beginning teacher of Filipino in California public schools.

For this item, which of the points on the scoring scale represents the level of response that would be achieved by this individual?

After panel members completed the practice set of items, the facilitator polled them regarding their item ratings; facilitated a discussion to review the concept of the “just-acceptable candidate;” discussed how to make the standard setting judgment; discussed how to review and consider the marker responses; and answered questions about the rating process.

Following the practice set, panel members began the same rating process with the actual items used on the November 4, 2006 test forms. In responding to the standard setting question, panel members were asked to refer to the score point descriptions that are appropriate for the type of item under consideration (i.e., the descriptions associated with a three-point scale or those associated with a four-point scale). They were also asked to refer to the marker responses for each score point for each assignment. NES analyzed the individual results from these item judgments for use in Round Two of the standard setting process.

Round Two Standard Setting Ratings

Round Two of the standard setting process moved the panel from providing ratings at the item level to ratings made at the subtest level. Panel members were asked to provide separate preliminary passing score recommendations for each subtest.

For Round Two, panel members were provided the following materials.

- subject matter requirements and subtest descriptions
- Round One Item Rating Summary Sheet, which provided the sum of the median rating for each item across all panel members, doubled to reflect the actual combined scores examinees would receive from two scorers. The sheet also provided the sum of each panel member’s Round One item ratings doubled to reflect the actual combined scores

examinees would receive from two scorers. These individual ratings were listed in descending order by score value.

- Round Two Subtest Standard Setting Recommendation Form

(NOTE: Results of individual panel members were provided by identification number only to maintain the confidentiality of each person's ratings.)

Working independently, and considering their own ratings from Round One and the results of the group's ratings, each panel member provided a single, holistic, Round Two passing score recommendation for each subtest, representing the total number of points at the subtest level that would, in their judgment, be earned by the just-acceptable candidate. To make this recommendation, panel members responded to the following question.

Imagine a hypothetical candidate who is just at the level of knowledge and skills important for effective job performance as a beginning teacher of Filipino in California public schools.

What is the total score for the constructed-response items on the subtest (out of the total number of score points for this subtest) that would be obtained by this individual?

Following this rating activity, NES collected and analyzed the panel members' recommendations and informed the panelists of the results.

Round Three Standard Setting Ratings

The goal of Round Three of the standard setting process was to produce a passing standard recommendation for each subtest.

For Round Three, panel members were provided the following materials.

- subject matter requirements and subtest descriptions
- Round Two Results Summary Sheets, which included the panel's computed median, and each panel member's Round Two item rating
- Round Three Subtest Standard Setting Recommendation Form
- Summary Statistics Report, which included for each subtest a set of analyses showing in tabular form the percent of examinees from the first test administration who would pass each subtest, given all possible passing scores.

These materials helped to facilitate consideration by panel members of their ratings, the nature of the examinee sample, the goal of Round Three, the purpose of the CSET program, and the concept of the just-acceptable candidate and measurement error.

The panel was cautioned about making judgments based on small numbers of examinees, and was advised that the examinees at the first test administration may or may not reflect the same proportions of all the types and capabilities of examinees in the population that will take the test in the future.

After further discussion, panel members were asked to independently recommend a passing standard for each subtest by responding to the following question.

Imagine a hypothetical candidate who is just at the level of knowledge and skills important for effective job performance as a beginning teacher of Filipino in California public schools.

What is the total score for the items on the subtest (out of the total number of score points) that would be obtained by this individual?

As the final step to the standard setting studies, each panel member was asked to complete independently a meeting evaluation form regarding the training provided and the task in general.

NES compiled the results of the standard setting panel for use in the determination of the staff-recommended passing standards presented in this report.

APPENDIX C

CSET STANDARDS SETTING CONSIDERATIONS

CSET Standards Setting Considerations

As described in the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999), the standard setting process is a key piece of validity evidence supporting a testing program.

Defining the minimum level of knowledge and skill required for licensure or certification is one of the most important and difficult tasks facing those responsible for credentialing. Verifying the appropriateness of the cut score or scores on the tests is a critical element in validity. The validity of the inference drawn from the test depends on whether the standard for passing makes a valid distinction between adequate and inadequate performance. Often, panels of experts are used to specify the level of performance that should be required. Standards must be high enough to protect the public, as well as the practitioner, but not so high as to be unreasonably limiting. Verifying the appropriateness of the cut score or scores on a test used for licensure or certification is a critical element of the validity of test results (p.157).

In making recommendations to the Commission on passing standards for the CSET: Filipino, staff considered the following factors and options that affect the standard setting process in determining the staff-recommended passing standards.

Subtest Scoring Model

The subtest scoring model used with CSET is a non-compensatory subtest model in which all subtests in a subject area must be passed independently. A subject matter advisory panel considered this model when determining the subtest structures of the Languages Other Than English examinations.

Professional Judgments

The recommended passing standards for the CSET are based upon the professional judgments provided by the members of the Subject matter advisory panels. Since these panel recommendations are criterion-referenced—based on expert judgment of the minimum required subject matter knowledge for beginning teachers—examinee performance data provides supplemental, though not necessary, information.

Standard Error of Measurement

Standard error of measurement is one way to express test reliability and addresses the imprecision of test data. Measurements are not perfectly reliable. In testing, for example, only one score from a single test administration is available for each examinee. An individual examinee's score may, or may not, be accurate. However, the standard error allows us to determine a range within which the examinee's score is likely to lie. Within reasonable limits, the standard error of measurement provides a safeguard against placing undue emphasis on a single numerical score. This is just one index of reliability, and should be applied to the standard setting process in combination with other test-specific characteristics.